

#8
BP
9-6-01 1638

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/548,971

DATE: 08/07/2001

TIME: 08:16:07

Input Set : A:\19452a-7.app

Output Set: N:\CRF3\08072001\I548971.raw

RECEIVED

SEP 04 2001

TECH CENTER 1600/2900

3 <110> APPLICANT: Liljegren, Sarah
 4 Yanofsky, Martin F.
 5 The Regents of the University of California
 7 <120> TITLE OF INVENTION: Control of Fruit Dehiscence in Arabidopsis by
 8 INDEHISCENT1 Genes
 10 <130> FILE REFERENCE: 19452A-000700US
 12 <140> CURRENT APPLICATION NUMBER: US 09/548,971
 13 <141> CURRENT FILING DATE: 2000-04-13
 15 <150> PRIOR APPLICATION NUMBER: US 60/090,649
 16 <151> PRIOR FILING DATE: 1998-06-25
 18 <150> PRIOR APPLICATION NUMBER: US 09/339,998
 19 <151> PRIOR FILING DATE: 1999-06-25
 21 <160> NUMBER OF SEQ ID NOS: 9
 23 <170> SOFTWARE: PatentIn Ver. 2.1
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 3856
 27 <212> TYPE: DNA
 28 <213> ORGANISM: Arabidopsis thaliana
 30 <220> FEATURE:
 31 <223> OTHER INFORMATION: INDEHISCENT1 (IND1) genomic
 33 <220> FEATURE:
 34 <221> NAME/KEY: CDS
 35 <222> LOCATION: (2765)..(3361)
 36 <223> OTHER INFORMATION: IND1
 38 <400> SEQUENCE: 1
 39 ctctagacca tctactatcc ggttggtgac ccttaaagct tttgaagact actagaataa 60
 41 tgcaaatacc atatgtccat atccatcctt ttcttttggt tgaactgaac attctaattt 120
 43 tgtaaaagaa aaaaccttat gttaatatca ccgtaggcaa aaaaaatatc tcatcatatt 180
 45 aaatttttat tataagatta tacattctct cgttgtaaga gttactccaa ttgcaagtgt 240
 47 tgtatttaact aataaaaagg acgaaagtag gaagccttata attaatgat gttgcatagt 300
 49 actggtatat tgttgatgaa tataacaagt atgaacatta atgcatgaaa cggggtat 360
 51 tgtcttgaac tcattaaagg caatgtgaaa agaagatgtg aggtctcatt ttgaaaattt 420
 53 atcttctagc tttgtcgatt ttaaatctat gaaatgaacg caacatatag aaatttcatt 480
 55 tggacaacga catttagacg gtatcttaat tagaccgatt aattagtaat atacttatat 540
 57 atataattag tggtgattat aagtttactt atccacttga gaatttaaac aatgggcaat 600
 59 accttaatgt cgaaagaagc cgtcccccact tcgtgtaatg agttatgggg gagagatcct 660
 61 gttaaatcgt caaataaaac aacttaagaa ctagaattg acaccaaaaa tcataaagag 720
 63 aacgttgaag aagtcattta tcgtatccag ctcatatttc cttagctaaga tcaaatcaag 780
 65 gccgttgaaa gggcttgtaa gaaaatgtcg aagaaaccgt ggggtttaga agaaagacaa 840
 67 gaaatagaag aacaatgatg ttaaatgtcc tatttttggt tataggagtt gtcaaaagag 900
 69 gagagagaga agaaaattag gtcaaaataa tgagcactaa aaatggagac atgtgttgag 960
 71 taactattac aagagcgact tatgcttcct tatggcaatg atatccaaac caaagtgcaa 1020
 73 cgctcccttt ttgccctaatt ttcgtaaagt ctctctcctt ctctgctcctt aggaaaaacc 1080
 75 ctagaattt aatcccttgt tcttgatcct gctttttgag taaccatgat tttgaccaca 1140
 77 cactatttct tctatctttt gtggtctata ggattttgct ttatatgtgt ttcttgtatt 1200
 79 gctccgtacg tacgtatacg aattttaatg gttataacaa ggtttatata aactagcaca 1260
 81 aatgagtcca tgaaatttgt tagcgaaaaa ggtagaataa tattgagtct ttaaaccggca 1320

Does Not Comply
Corrected Diskette Needed

PP-315

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/548,971

DATE: 08/07/2001

TIME: 08:16:07

Input Set : A:\19452a-7.app

Output Set: N:\CRF3\08072001\I548971.raw

```

83 atatatataa ttttgctgca aaacttagct ttaatcatga tctaatagata ttttctttaa 1380
85 tttcctttgc caaattaatc acatgcacgg atttttggca agttatgtgt cgaattcttc 1440
87 cattcacaca aactaaact taattagaac tctaggaaat attttaaaat gacaacttta 1500
89 tcgaaaaaaa tttagttatg aaaacaattc cagaattaaa catgagctat ataatttaag 1560
91 ataaaatgaa gtaatatgta tatgtatgta ataacatata tgattgcggt aaaaaaaac 1620
93 atatctgatt aaattgttca tgcaggccca tgtcactatg atgtcatcac gtttttattt 1680
95 tcacaataac taatatatat tcaaaaaaat agttttgtca gattaaattt tttttggtgg 1740
97 tcagctttct ccaacctact aaactagttt ggaatgttct cttctttatt tttcttttct 1800
99 ttgatttctt atgtttttta tttatggaat ttaagacgg attgtttagg tcgtttctct 1860
101 cttttcttgt tttctaaagt tacttttgta aactcatctc ctcccaatta gacagtcaat 1920
103 catatagtta tcttttaata tatgtctagt tgataaaaaa aatgaaaaaa tactggtggt 1980
105 agttctacta atgtttgtgt aaaaaatctg atattatgaa tctaataaat ttctttgatc 2040
107 gtataatgtg gggttaaattt agtaattttt tacataaata agaactgtaa tgttgatgta 2100
109 tattggggaa tcagtatatg agcttgggta actatacttc tggaaatact tgaagattta 2160
111 actatttgca aaattataat ttagtcccga aaaatacaga cgacgggaca cgacaacata 2220
113 taagcaggtt tgaatcttgg aaaattttgt atacataacc tatataaata ctaatgttct 2280
115 gggtgggttc aaaagccttt tcaaaagtcc cattttttta attcaaggac attttacata 2340
117 ggaaataagt tgagtcataa aaaataatgg ttattttgta aggttttttt tttgattaaa 2400
119 acgcacatat taagaagtta gtttttttct actaccaaat atcaattaat ttaaaaccat 2460
121 gcaaccattc ataaaacaat actattaaag aatataaata atcacaaat attaaatata 2520
123 cttaaaattt acatataaat ttacaaaaca tctaattaat tgaaacagaa aggaaaagg 2580
125 aaaatatata ataaaatgag acatatatcc tataaaaaaa aatgaggca tatgaagtaa 2640
127 ataataagag acatgcattg aagcattcgg ttaattaatc gagtcaaaga tatatatcag 2700
129 taaatacata tgtgtatatg tctggaaaaa gaatatatat attgagaaat aagaaaagat 2760
131 gaaa atg gaa aat ggt atg tat aaa aag aaa gga gtg tgc gac tct tgt 2809
132 Met Glu Asn Gly Met Tyr Lys Lys Lys Gly Val Cys Asp Ser Cys
133 1 5 10 15
135 gtc tcg tcc aaa agc aga tcc aac cac agc ccc aaa aga agc atg atg 2857
136 Val Ser Ser Lys Ser Arg Ser Asn His Ser Pro Lys Arg Ser Met Met
137 20 25 30
139 gag cct cag cct cac cat ctc ctc atg gat tgg aac aaa gct aat gat 2905
140 Glu Pro Gln Pro His His Leu Leu Met Asp Trp Asn Lys Ala Asn Asp
141 35 40 45
143 ctt ctc aca caa gaa cac gca gct ttt ctc aat gat cct cac cat ctc 2953
144 Leu Leu Thr Gln Glu His Ala Ala Phe Leu Asn Asp Pro His His Leu
145 50 55 60
147 atg tta gat cca cct ccc gaa acc cta att cac ttg gac gaa gac gaa 3001
148 Met Leu Asp Pro Pro Pro Glu Thr Leu Ile His Leu Asp Glu Asp Glu
149 65 70 75
151 gag tac gat gaa gac atg gat gcg atg aag gag atg cag tac atg atc 3049
152 Glu Tyr Asp Glu Asp Met Asp Ala Met Lys Glu Met Gln Tyr Met Ile
153 80 85 90 95
155 gcc gtc atg cag ccc gta gac atc gac cct gcc acg gtc cct aag ccg 3097
156 Ala Val Met Gln Pro Val Asp Ile Asp Pro Ala Thr Val Pro Lys Pro
157 100 105 110
159 aac cgc cgt aac gta agg ata agc gac gat cct cag acg gtg gtt gct 3145
160 Asn Arg Arg Asn Val Arg Ile Ser Asp Asp Pro Gln Thr Val Val Ala
161 115 120 125
163 cgt cgg cgt cgg gaa agg atc agc gag aag atc cga att ctc aag agg 3193

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/548,971

DATE: 08/07/2001

TIME: 08:16:08

Input Set : A:\19452a-7.app

Output Set: N:\CRF3\08072001\I548971.raw

```

164 Arg Arg Arg Arg Glu Arg Ile Ser Glu Lys Ile Arg Ile Leu Lys Arg
165      130      135      140
167 atc gtg cct ggt ggt gcg aag atg gac aca gct tcc atg ctc gac gaa 3241
168 Ile Val Pro Gly Gly Ala Lys Met Asp Thr Ala Ser Met Leu Asp Glu
169      145      150      155
171 gcc ata cgt tac acc aag ttc ttg aaa cgg cag gtg agg att ctt cag 3289
172 Ala Ile Arg Tyr Thr Lys Phe Leu Lys Arg Gln Val Arg Ile Leu Gln
173 160      165      170      175
175 cct cac tct cag att gga gct cct atg gct aac ccc tct tac ctt tgt 3337
176 Pro His Ser Gln Ile Gly Ala Pro Met Ala Asn Pro Ser Tyr Leu Cys
177      180      185      190
179 tat tac cac aac tcc caa ccc tga tgaactacac agaagctcgc tagctagaca 3391
180 Tyr Tyr His Asn Ser Gln Pro
181      195
183 tttggtgtca tcctctcaac ctttttcatg ttgatattat atatatagat gcataaagat 3451
185 tcgatccaag attgtatggg tgttttaata ttattattct aagatatatg atgtacaatt 3511
187 gtgtaccaag tttctttatc ttgatattcat atgcataaat aattggtgaa taaaaagaag 3571
189 atattgattg taaacaaaaa aaagaagata ttgattgtta attagggttt gatcattctg 3631
191 tatgaaagct ttggcctgca aattaatttt cgatatatat atatataat ggagaatata 3691
193 tatcaaatac ttttttaatt tgactataat ttgtatcaat tatctgaatc tgatgagtgt 3751
195 aggttatata tggattagca aaaaagaaaa caaccattat tacgcaccta cattaaaaat 3811
197 catccaccaa agaagaaacc atcctcaaga gggttccctc tagag 3856
200 <210> SEQ ID NO: 2
201 <211> LENGTH: 198
202 <212> TYPE: PRT
203 <213> ORGANISM: Arabidopsis thaliana ✓
204 <223> OTHER INFORMATION: INDEHISCENT1 (IND1) protein need field 220
206 <400> SEQUENCE: 2
207 Met Glu Asn Gly Met Tyr Lys Lys Lys Gly Val Cys Asp Ser Cys Val
208 1 5 10 15
209 Ser Ser Lys Ser Arg Ser Asn His Ser Pro Lys Arg Ser Met Met Glu
210 20 25 30
211 Pro Gln Pro His His Leu Leu Met Asp Trp Asn Lys Ala Asn Asp Leu
212 35 40 45
213 Leu Thr Gln Glu His Ala Ala Phe Leu Asn Asp Pro His His Leu Met
214 50 55 60
215 Leu Asp Pro Pro Pro Glu Thr Leu Ile His Leu Asp Glu Asp Glu Glu
216 65 70 75 80
217 Tyr Asp Glu Asp Met Asp Ala Met Lys Glu Met Gln Tyr Met Ile Ala
218 85 90 95
219 Val Met Gln Pro Val Asp Ile Asp Pro Ala Thr Val Pro Lys Pro Asn
220 100 105 110
221 Arg Arg Asn Val Arg Ile Ser Asp Asp Pro Gln Thr Val Val Ala Arg
222 115 120 125
223 Arg Arg Arg Glu Arg Ile Ser Glu Lys Ile Arg Ile Leu Lys Arg Ile
224 130 135 140
225 Val Pro Gly Gly Ala Lys Met Asp Thr Ala Ser Met Leu Asp Glu Ala
226 145 150 155 160
227 Ile Arg Tyr Thr Lys Phe Leu Lys Arg Gln Val Arg Ile Leu Gln Pro

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/548,971

DATE: 08/07/2001

TIME: 08:16:08

Input Set : A:\19452a-7.app

Output Set: N:\CRF3\08072001\I548971.raw

```

228           165           170           175
229 His Ser Gln Ile Gly Ala Pro Met Ala Asn Pro Ser Tyr Leu Cys Tyr
230           180           185           190
231 Tyr His Asn Ser Gln Pro
232           195
235 <210> SEQ ID NO: 3
236 <211> LENGTH: 2765
237 <212> TYPE: DNA
238 <213> ORGANISM: Arabidopsis thaliana
240 <220> FEATURE:
241 <223> OTHER INFORMATION: 5' untranslated region IND1 promoter OK
243 <400> SEQUENCE: 3
244 ctctagacca tctactatcc ggttggtgac ccttaaagct tttgaagact actagaataa 60
245 tgcaaatacc atatgtccat atccatcctt ttcttttggt tgaactgaac attctaattt 120
246 tgtaaaagaa aaaaccttat gttaatatca ccgtaggcaa aaaaaatata tcatcatatt 180
247 aaatttttat tataagatta tacattctct cgttgtaaga gttactccaa ttgcaagtgt 240
248 tgtattaact aataaaaagg acgaaagtag gaagcttata attaattgat gttgcatagt 300
249 actggtatat tggtgatgaa tataacaagt atgaacatta atgcatgaaa cgggggtattt 360
250 tgtcttgaac tcattaaagg caatgtgaaa agaagatgtg aggtctcatt ttgaaaattt 420
251 atcttctagc tttgtcgatt ttaaatctat gaaatgaacg caacatatag aaatttcatg 480
252 tggacaacga catttagacg gtatcttaat tagaccgatt aattagtaat atacttatat 540
253 atataattag tgggtgattat aagtttactt atccacttga gaatttaaac aatgggcaat 600
254 accttaatgt cgaaagaagc cgtccccact tcgtgtaatg agttatgggg gagagatcct 660
255 gttaaatcgt caaataaaaac aacttaagaa ctagaaattg acaccaaaaa tcataaagag 720
256 aacgttgaag aagtcattta tcgtatccag ctcatatttc ctagctaaga tcaaatcaag 780
257 gccgttgaaa gggcttgtaa gaaaatgtcg aagaaaccgt ggggtttaga agaaagacaa 840
258 gaaatagaag aacaatgatg ttaaatggcc tattttgggt tataggagtt gtcaaaagag 900
259 gagagagaga agaaaattag gtcaaaataa tgagcactaa aaatggagac atgtgttgag 960
260 taactattac aagagcgact tatgcttcct tatggcaatg atatccaaac caaagtgcaa 1020
261 cgctcctttt ttgccctaata ttcgtaaagt ctctctcctt ctctcgtcctt aggaaaaacc 1080
262 ctagaaattt aatcccttgt tcttgatcct gctttttgag taaccatgat tttgaccaca 1140
263 cactatttct tctatctttt gtggtctata ggattttgct ttatatgtgt ttcttgtatt 1200
264 gctccgtacg tacgtatacg aattttaaag gttataacaa ggtttatata aactagcaca 1260
265 aatgagtcca tgaaatttgt tagcgaaaaa ggtagaaata tattgagtct ttaaacggca 1320
266 atatatataa ttttgctgca aaacttagct ttaatcatga tctaatagata ttttctttaa 1380
267 tttcctttgc caaattaatc acatgcacgg atttttggca agttatgtgt cgaattcttc 1440
268 cattcacaca acactaaact taattagaac tctaggaaat attttaaaat gacaacttta 1500
269 tcgaaaaaaa tttagttatg aaaacaattc cagaattaaa catgagctat ataatttaag 1560
270 ataaaatgaa gtaatatgta tatgtatgta ataacatata tgattgcggg aaaaaaaaac 1620
271 atatctgatt aaattgttca tgcaggccca tgtcactatg atgtcatcac gtttttattt 1680
272 tcacaataac taatatatat tcaaaaaaat agttttgtca gattaaattt tttttggtgg 1740
273 tcagctttct ccaacctact aaactagttt ggaatgttct cttctttatt tttctttttc 1800
274 ttgatttctt atgtttttta tttatggaat tttaagacgg attgtttagg tcgtttctct 1860
275 cttttcttgt tttctaaagt tacttttgta aactcatctc ctcccaatta gacagtcaat 1920
276 catatagtta tcttttaata tatgtctagt tgataaaaaa aatgaaaaaa tactggtggg 1980
277 agttctacta atgtttgtgt aaaaaatctg atattatgaa tctaataaat ttctttgatc 2040
278 gtataatgtg gggttaaattt agtaattttt tacataaata agaactgtaa tgttgatgta 2100
279 tattggggaa tcagtatatt agcttgggta actatacttc tggaaatact tgaagattta 2160
280 actatttgca aaattataat ttagtcccga aaaatacaga cgacgggaca cgacaacata 2220

```


RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/548,971

DATE: 08/07/2001
TIME: 08:16:08

Input Set : A:\19452a-7.app
Output Set: N:\CRF3\08072001\I548971.raw

```

281 taagcagggtt tgaatcttgg aaaattttgt atacataacc tatataaata ctaatgttct 2280
282 gggttggttc aaaagccttt tcaaaagttc cattttttta attcaaggac attttacata 2340
283 ggaaataagt tgagtcataa aaaataatgg ttatttttga aggttttttt tttgattaaa 2400
284 acgcacatat taagaagtta gttttttttc actaccaa atcaattaat ttaaaacat 2460
285 gcaaccattc ataaaacaat actattaaag aatataaata atcacaaaat attaaataca 2520
286 cttaaaattt acatataaat ttacaaaaca tctaattaat tgaaacagaa aggaaaagg 2580
287 aaaatatatc ataaaatgag acatatatcc tataaaaaaa aaatgaggca tatgaagtaa 2640
288 ataataagag acatgcatgt aagcattcgg ttaattaatc gagtcaaaga tatatatcag 2700
289 taaatacata tgtgtatatt tctggaaaaa gaatatatat attgagaaat aagaaaagat 2760
290 gaaaa

```

293 <210> SEQ ID NO: 4

294 <211> LENGTH: 496

295 <212> TYPE: DNA

296 <213> ORGANISM: Arabidopsis thaliana

298 <220> FEATURE:

299 <223> OTHER INFORMATION: 3' untranslated region IND1 promoter

301 <400> SEQUENCE: 4

```

302 atgaactaca cagaagctcg ctagctagac atttggtgtc atcctctcaa cttttttcat 60
303 gttgatatat tatatataga tgcataaaga ttcatccaa gattgtatgg gtgtttta 120
304 attattatc taagatatat gatgtacaat tgtgtacca gtttctttat cttgatata 180
305 tatgcataaa taattggtga ataaaagaa gatattgatt gtaaacaaaa aaaagaagat 240
306 attgattggt aattagggtt tgatcattct gtatgaaagc tttggcctgc aaattaattt 300
307 tcatatata tatatatata tggagaatat atatcaata cttttttta ttgactataa 360
308 tttgtatcaa ttatctgaat ctgatgagtg taggttatat atggattagc aaaaaagaa 420
309 acaaccatta ttacgcacct acattaaaaa tcatccacca aagaagaaac catcctcaag 480
310 agggttccct ctagag

```

313 <210> SEQ ID NO: 5

314 <211> LENGTH: 5622

315 <212> TYPE: DNA

316 <213> ORGANISM: Arabidopsis thaliana

318 <220> FEATURE:

319 <223> OTHER INFORMATION: SHATTERPROOF1 (SHP1) genomic

321 <220> FEATURE:

322 <221> NAME/KEY: modified_base

323 <222> LOCATION: (935)..(941)

324 <223> OTHER INFORMATION: n = any nucleotide

326 <400> SEQUENCE: 5

```

327 agatctgcaa cagtgaagaa agaaaacaaa atggacttga agaggttttg acaatgccag 60
328 agataatgct tattccctaa tatgttgcca gccagtgtc aaattggctt tttaaata 120
329 gatttctgta tcagtgtgca tatttggtga tccaacgtat tcatcatcaa gttctcaag 180
330 ttgctttcag tgcaattcta attcacacgt ttaactttta catgcatgtc attataatta 240
331 cttcttctact aagacacaat acggcaaacc tttcagatta tattaatctc cataaatgaa 300
332 ataattaacc tcataatcaa gattcaatgt ttctaaatat atatggacaa aatttacacg 360
333 gaagattaga tacgtatatt agtagattta gtctttcggt tgtgcgataa gattaaccac 420
334 ctcatagata gtaatatcat tgtcaaatc ctctcggttt agtcgctaaa ttgtatcttt 480
335 tttaagccta aaagtagtgt attcgcatat gacttatcgt cctaactttt tttttaatta 540
336 acaaaaaaat cgaaaagaaa ataattctgt aaatatTTTT taagtactcc attagttta 600
337 gtttctatTTT aaaaaatgct tgaaatttga cagttatgtt caacaatttt gaatcatgag 660
338 cgatgtctag atactcagaa tttaatcaag atgtcttata aaatttggtg tcaactcgagg 720

```

Use of n and/or Xaa has been detected in the Sequence Listing.
Review the Sequence Listing to insure a corresponding
explanation is presented in the <220> to <223> fields of
each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/548,971

DATE: 08/07/2001

TIME: 08:16:09

Input Set : A:\19452a-7.app

Output Set: N:\CRF3\08072001\I548971.raw

L:342 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5